

## Durham Research Online

---

### Deposited in DRO:

24 September 2019

### Version of attached file:

Published Version

### Peer-review status of attached file:

Peer-reviewed

### Citation for published item:

Leo, Matteo and Arnold, Christian and Li, Baojiu (2019) 'A high-redshift test of gravity using enhanced growth of small structures probed by the neutral hydrogen distribution.', *Physical review D.*, 100 (6). 064044.

### Further information on publisher's website:

<https://link.aps.org/doi/10.1103/PhysRevD.100.064044>

### Publisher's copyright statement:

Reprinted with permission from the American Physical Society: Leo, Matteo, Arnold, Christian Li, Baojiu (2019). A high-redshift test of gravity using enhanced growth of small structures probed by the neutral hydrogen distribution. *Physical Review D* 100: 064044. © 2019 by the American Physical Society. Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or part, without prior written permission from the American Physical Society.

### Additional information:

## Use policy

---

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

# High-redshift test of gravity using enhanced growth of small structures probed by the neutral hydrogen distribution

Matteo Leo<sup>1,2,\*</sup>, Christian Arnold,<sup>2</sup> and Baojiu Li<sup>2</sup>

<sup>1</sup>*Institute for Particle Physics Phenomenology, Department of Physics, Durham University, Durham DH1 3LE, United Kingdom*

<sup>2</sup>*Institute for Computational Cosmology, Department of Physics, Durham University, Durham DH1 3LE, United Kingdom*



(Received 16 January 2019; published 23 September 2019)

Future 21-cm intensity mapping surveys such as SKA can provide precise information on the spatial distribution of the neutral hydrogen (HI) in the postreionization epoch. This information will allow us to test the standard  $\Lambda$  cold dark matter paradigm and with that the nature of gravity. In this work, we employ the SHYBONE simulations, which model galaxy formation in  $f(R)$  modified gravity using the IllustrisTNG model, to study the effects of modified gravity on HI abundance and power spectra. We find that the enhanced growth low-mass dark matter halos experience in  $f(R)$  gravity at high redshifts alters the HI power spectrum and can be observable through 21-cm intensity mapping. Our results suggest that the HI power spectrum is suppressed by  $\sim 13\%$  on scales  $k \lesssim 2h \text{ Mpc}^{-1}$  at  $z = 2$  for F6, a  $f(R)$  model which passes most observational constraints. We show that this suppression can be detectable by SKA1-MID with 1000 hours of exposure time, making HI clustering a novel test of gravity at high redshift.

DOI: [10.1103/PhysRevD.100.064044](https://doi.org/10.1103/PhysRevD.100.064044)

## I. INTRODUCTION

Our standard model of cosmology—the  $\Lambda$  cold dark matter model ( $\Lambda$ CDM)—has proved very successful in describing almost all currently available observational data of the Universe. Its underlying theory of gravity, Einstein’s general relativity (GR), has been tested to remarkably high precision on small scales [1]. In recent years, in the wake of high-precision astronomical observations, tests of GR on cosmological scales have become possible and commonplace as well [2], although until now these tests have primarily focused on comparatively high-mass objects and low redshifts (e.g., [3–12]). Because of the screening mechanisms which many alternatives to GR employ to pass the stringent Solar System tests, these objects are less suited to distinguish GR from alternative, *modified gravity* (MG) theories with screening mechanisms.

As a representative example, we consider a particular one of these MG models in this paper— $f(R)$  gravity [13], though we expect our conclusions to hold at least qualitatively for general thin-shell screening [14] models.  $f(R)$  gravity is a generalization of GR which alters cosmic structure formation through a factor-of-4/3 enhanced gravitational force. We adopt the popular variant proposed in Ref. [15], which, with certain choices of the free parameters of the model, can produce a cosmic expansion history very close to that of the  $\Lambda$ CDM paradigm.

The model employs the so-called *chameleon screening* mechanism [16,17] to ensure that the modifications to standard gravity are suppressed and GR-like behavior is recovered in high-density regions like the Solar System. The model considered here has been widely studied using numerical simulations, e.g., [9,18–24]. In this work we consider two instances of it [15]: F6 and F5, with its model parameter  $f_{R0}$  equal to  $-10^{-5}$  and  $-10^{-6}$ , respectively (see Sec. II A for more details).

Although the constraints from our local environment are very tight, as mentioned above, the previous constraints from cosmological scales are much weaker since the objects used in these tests are generally more massive and well screened. One way to overcome this limitation is to study low-mass objects which are less likely to be screened and hence experience larger deviations from GR. However, a major challenge to this approach is the difficulty to accurately detect, resolve, and trace such small objects in observations, even at low redshifts.

In this paper, we propose a novel test of gravity at intermediate scales and high redshifts ( $z \geq 2$ ), using the distribution of neutral hydrogen (HI) in our Universe, which is observable in 21-cm experiments (some current and future instruments of this kind include SKA [25], MeerKAT [26], LOFAR [27], CHIME [28], and BINGO [29]). 21-cm intensity mapping can be used to trace the underlying distribution of matter [30–33] and with that the low-mass halos in the Universe (as suggested in [34]). In order to determine how possible deviations from GR

\*matteo.leo@durham.ac.uk

would affect the HI distribution, we employ the SHYBONE simulations, a set of full-physics hydrodynamical simulations of  $f(R)$  modified gravity. Comparing the  $f(R)$  simulations to their  $\Lambda$ CDM counterpart allows us to quantify the size of the MG effects on HI observables (such as the overall neutral hydrogen abundance and the HI power spectrum), and to assess if these effects can be observable with future 21-cm intensity mapping experiments. HI clustering has been proposed as a probe for a number of nonstandard cosmological models, e.g., massive neutrinos, warm DM, dark energy, and modified gravity [35–39], but this study reveals new features, thanks to the high resolution of our simulations.

This paper is structured as follows. In Sec. II we briefly introduce the two instances of Hu-Sawicki (HS)  $f(R)$  gravity [15] used in our investigation and the suite of hydrodynamical simulations employed to quantify the abundance and clustering of HI. In Sec. III, we show and discuss the main results of this paper, including the overall neutral hydrogen density (III A), the HI abundance in halos (III B), and the HI power spectra in both real and redshift space (III C). Additional tests to explain the physics behind our results are performed in Sec. III D and observational forecasts for a future 21-cm intensity mapping experiment are discussed in Sec. III E. In Sec. III F, we comment on the dependence of our results on the galaxy formation model employed in our simulations. Finally, we conclude our findings in Sec. IV.

## II. THEORETICAL MODELS AND SIMULATIONS

### A. $f(R)$ gravity

$f(R)$  gravity is a popular class of MG models that is obtained by adding a scalar function  $f(R)$  to the Ricci scalar  $R$  in the standard Einstein-Hilbert action [13] of general relativity. With an appropriate choice of the functional form and parameters of  $f(R)$ , the theory can mimic the late time expansion history of a  $\Lambda$ CDM universe without explicitly having a cosmological constant  $\Lambda$  (the accelerated expansion in these theories is achieved via some form of quintessence/cosmological constant and is not due to the modification of gravity itself [40–42]).

The action for the  $f(R)$  gravity can be written as

$$S = \int d^4x \sqrt{-g} \left[ \frac{R + f(R)}{16\pi G} + \mathcal{L}_m \right], \quad (1)$$

where  $G$  is the gravitational constant,  $g$  is the determinant of the metric,  $g_{\mu\nu}$ , and  $\mathcal{L}_m$  is the standard matter/radiation Lagrangian density. The simulations considered here employ the weak-field and quasistatic limit (see [43] for more details on the validity of these approximations), so that the equations of motion obtained by varying the action in Eq. (1) can be simplified to a (modified) Poisson

equation plus an equation for the scalar degree of freedom,  $f_R \equiv df(R)/dR$  (the so-called *scalar* field),

$$\nabla^2 \Phi = \frac{16\pi G}{3} \delta\rho - \frac{1}{6} \delta R, \quad (2)$$

$$\nabla^2 f_R = \frac{1}{3} (\delta R - 8\pi G \delta\rho), \quad (3)$$

where  $\delta\rho \equiv \rho - \bar{\rho}$  and  $\delta R \equiv R - \bar{R}$  are the matter density perturbation and the Ricci scalar perturbation, respectively (and  $\bar{\rho}$  and  $\bar{R}$  are their background values).

The HS variant of the theory [15] uses

$$f(R) = -m^2 \frac{c_1 \left(\frac{R}{m^2}\right)^n}{c_2 \left(\frac{R}{m^2}\right)^n + 1}, \quad (4)$$

where  $m$  is a new mass scale of the model,  $m^2 \equiv \Omega_m H_0^2$ ,  $H_0$  is the Hubble constant,  $\Omega_m$  is the total nonrelativistic matter energy density at present time in units of the present-day critical energy density of the Universe,  $\rho_{c0} \equiv 3H_0^2/8\pi G$ , and  $c_1, c_2, n$  are model parameters. We choose  $n = 1$  hereafter for simplicity. Furthermore, if we tune the parameters  $c_1$  and  $c_2$  such that

$$\frac{c_1}{c_2} = 6 \frac{\Omega_\Lambda}{\Omega_m} \quad \text{and} \quad \frac{c_2 |R|}{m^2} \gg 1, \quad (5)$$

the model leads to a cosmic expansion history which is very close to that of a  $\Lambda$ CDM universe [15].  $\Omega_\Lambda$  in the above equation represents the cosmological constant energy density in units of  $\rho_{c0}$  for the  $\Lambda$ CDM universe; in the case of  $f(R)$  gravity,  $\Omega_\Lambda$  still enters in the theory as a parameter. If one further assumes that the Universe is spatially flat, then  $\Omega_\Lambda$  is simply given by  $\Omega_\Lambda = 1 - \Omega_m$ . This is our default assumption in this work.

The scalar field  $f_R$  in this model can be approximated as

$$f_R \equiv \frac{df(R)}{dR} \approx -\frac{c_1}{c_2^2} \left(\frac{m^2}{R}\right)^2, \quad (6)$$

and its background evolution can be expressed in terms of the background Ricci scalar  $\bar{R}$ ,

$$\bar{f}_R(a) = \bar{f}_{R0} \left[ \frac{\bar{R}_0}{\bar{R}(a)} \right]^2, \quad (7)$$

where  $\bar{R}_0$  is the value of the Ricci scalar today and

$$\bar{R}(a) = 3m^2 \left[ a^{-3} + 4 \frac{\Omega_\Lambda}{\Omega_m} \right]. \quad (8)$$

The theory is therefore fully specified by  $\Omega_m$  and the present-day value of the background scalar field,  $\bar{f}_{R0}$ .

In order to satisfy the stringent constraints on the possible deviations from standard gravity in our local

environment [1], the theory employs the above-mentioned chameleon screening mechanism [16,17] to suppress modifications to gravity and restore GR in high-density regions. The chameleon screening has been described in great detail in the literature and thus we will not discuss it further here, but instead simply mention that it becomes effective when  $f_R$  becomes close to zero, such that  $\delta R \approx 8\pi G\delta\rho$  according to Eq. (3), and then Eq. (2) reduces to the standard Poisson equation in Newtonian gravity. The screening is more likely to take place at earlier times when matter density is high and the background value of the scalar field,  $|\bar{f}_R|$ , is small. At a given time, this mechanism screens regions where their density is high and therefore the Newtonian potential is deep. The transition between screened and unscreened regimes depends on the choice of  $\bar{f}_{R0}$ . In  $f(R)$  gravity, the speed of the gravitational wave is equal to the speed of light and the model passes recent constraints from gravitational wave observations [44], making it one of the most promising alternatives to GR.

As mentioned in the Introduction, in this work we will focus on the F5 and F6 instances of HS  $f(R)$  gravity, for which  $\bar{f}_{R0}$  is equal to  $-10^{-5}$  and  $-10^{-6}$ , respectively. In F5, the effects of MG are stronger than in F6 and this model is now ruled out by observational constraints [12] (see [45] for a review about the recent constraints on chameleon gravity). However, it is used here as a toy model to assess the effects of a stronger deviation from GR on the HI distribution. The F6 model is consistent with most cosmological observations.

An important characteristic of the chameleon screening is that this mechanism becomes inefficient for small structures at high redshift, while more massive objects and denser environments become unscreened at later times (lower redshift). At high redshift, low-mass halos are already unscreened and affected by modified gravity. As a consequence, by observing such small structures we can, in principle, place constraints on the  $f(R)$  deviations from GR. As we will see in the next sections, 21-cm intensity mapping is sensitive to the abundance of halos down to  $10^9 M_\odot$ , making it a very promising probe of differences at the low-mass end of the halo mass function, without the need to resolve individual halos.

### B. Full-physics simulations in MG

In order to quantify how modifications to gravity affect the 21-cm signal, we analyze the SHYBONE simulations [46], a set of high-resolution full-physics hydrodynamical simulations of HS  $f(R)$  gravity, carried out with the moving mesh simulation code AREPO [47]. The suite includes two subsets of simulations: a large-box set with a box size of  $L = 62h^{-1}$  Mpc (S62 hereafter) and a small-box set for which  $L = 25h^{-1}$  Mpc (S25 hereafter), both with roughly  $2 \times 512^3$  resolution elements [ $h$  is the

dimensionless Hubble constant, given by  $h \equiv H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1})$ ]. The S62 simulations have a mass resolution of  $m_{\text{DM}} = 1.3 \times 10^8 h^{-1} M_\odot$  for DM particles and roughly  $m_{\text{gas}} = 2.5 \times 10^7 h^{-1} M_\odot$  for gas cells, and they have been run for GR, F6, and F5 up to  $z = 0$ . The S25 simulations have a mass resolution of  $m_{\text{DM}} = 8.4 \times 10^6 h^{-1} M_\odot$  and  $m_{\text{gas}} = 1.6 \times 10^6 h^{-1} M_\odot$  and have been run for the same three models, up to  $z = 0$  for GR and F6 and up to  $z = 1$  for F5 (the enhanced gravitational interactions in the F5 model considerably increase the computational cost of the simulations compared to their GR counterpart). S25 features a higher resolution, but its smaller box means that we inevitably lose some information of large-scale modes and massive halos (see the discussion in the next section). The S62 suite features also DM-only (DMO hereafter) counterparts for all the runs, which are used to compare the halo mass function from full-physics and DMO simulations below. All simulations adopt the Planck 2016 [48] cosmology with  $\Omega_m = 0.3089$ ,  $\Omega_B = 0.0486$ ,  $\Omega_\Lambda = 0.6911$ ,  $h = 0.6774$ ,  $\sigma_8 = 0.8159$ , and  $n_s = 0.9667$ , where  $\Omega_B$  is the present-day baryon density parameter,  $\sigma_8$  is the root-mean-squared matter density fluctuation over spherical regions with radius  $8h^{-1}$  Mpc at  $z = 0$ , and  $n_s$  is the index of the primordial power spectrum.

The full-physics simulations use the IllustrisTNG hydrodynamical model [49–57], incorporating a prescription of star and black hole formation and feedback, gas cooling, galactic winds, and magnetohydrodynamics on a moving Voronoi mesh [53,57]. The equations for  $f(R)$  gravity are solved to full nonlinearity in the Newtonian limit by the modified gravity solver in the code [46], fully capturing the effects of the chameleon screening.

To calculate the neutral hydrogen fraction in each Voronoi cell, we follow the prescription in Sec. 2.2 in [34]. For non-star-forming gas, we use the neutral hydrogen fraction calculated on the fly in the simulations, while for star-forming gas we postprocess the outputs, recalculating the neutral hydrogen fraction in each cell assuming a temperature of  $T = 10^4$  K and following the approach in [58] to take into account self-shielding corrections. The postprocessing gives the total fraction of hydrogen that is nonionized: atomic (HI) and molecular hydrogen ( $\text{H}_2$ ). Because we are solely interested in HI, we calculate and subtract the fraction of  $\text{H}_2$  for each Voronoi cell as in [34].

## III. RESULTS

### A. Overall neutral hydrogen density

In Fig. 1 we show the overall neutral hydrogen density measured from the S62 (in the range  $0 \leq z \leq 3$  for all models) and S25 (in the range  $0 \leq z \leq 5$  for GR and F6 and  $1 \leq z \leq 5$  for F5) simulations. We follow the common definition for the overall HI abundance,  $\Omega_{\text{HI}}(z) = \bar{\rho}_{\text{HI}}(z)/\rho_{\text{c0}}$ , where  $\bar{\rho}_{\text{HI}}(z)$  is the mean HI density in our



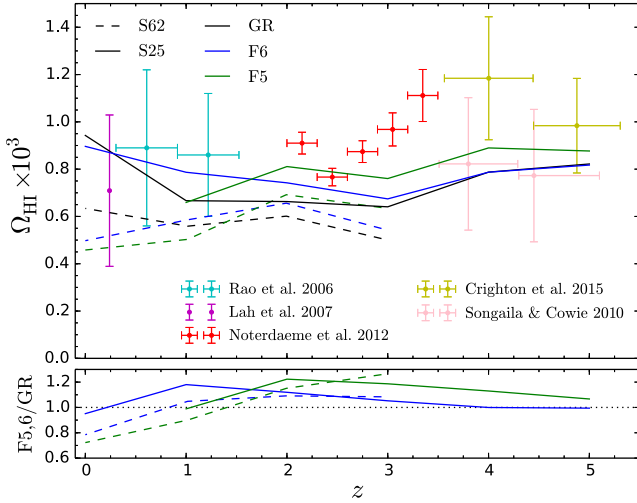


FIG. 1. Top panel: Overall HI abundance,  $\Omega_{\text{HI}}(z) = \bar{\rho}_{\text{HI}}(z)/\rho_{c0}$ , where  $\bar{\rho}_{\text{HI}}(z)$  is the mean HI density, from GR (black), F6 (blue), and F5 (green), compared with observationally measured values (symbols). Solid lines refer to S25 simulations, while dashed lines refer to S62 simulations. Bottom panel: The relative differences of the simulation predictions from F6 (blue) and F5 (green) with respect to GR.

simulations at a given redshift  $z$  and  $\rho_{c0}$  is the present-day critical density as defined above. We also show a selection of observational data for the HI abundance at different redshifts from [59–63].

First, we note that in Fig. 1 the HI abundance (for each model) predicted by S25 is higher than that measured from the low-resolution S62. A similar effect was found in [34] comparing the low- and high-resolution TNG simulations. This discrepancy between simulations at different resolution can be understood as follows. The neutral hydrogen in the postreionization epoch is concentrated in halos, where shielding effects screen them from ionization. It was shown (see, e.g., [34]; but see also the next section) that there is a significant amount of HI in halos with masses as low as  $10^9 M_\odot$  at  $z \leq 5$ . This implies that resolving halos of this mass in simulations is essential to measure the HI abundance accurately. However, because of its lower resolution, S62 does not fully resolve halos with masses  $< 10^{10} M_\odot$ , and therefore predicts a lower value for the HI abundance at all redshifts considered in this analysis. On the other hand, S25 can resolve halos down to  $6 \times 10^8 M_\odot$ , and therefore produces more reliable results of the HI abundance. However, we note that because of their small-box size, we do not have a statistically robust sample of halos with masses  $\gtrsim 10^{12} M_\odot$  in S25. This explains why our high-resolution simulations predict slightly lower HI abundance at  $z \leq 5$  compared to that measured in [34] for GR using the TNG-100 simulation (performed in a box of comoving length  $L = 75h^{-1}$  Mpc at the same resolution as S25).

Considering the differences between  $f(R)$  gravity and GR, in S25 the ratios of  $\Omega_{\text{HI}}(z)$  with respect to GR (solid

lines in the lower panel of Fig. 1) show a similar trend in F6 and F5. Indeed, in both models,  $\Omega_{\text{HI}}$  is similar to GR at high redshifts, larger than GR at intermediate redshifts (with an enhanced peak at  $z = 1$  and  $z = 2$  for F6 and F5, respectively) and falls below the GR values for  $z < 1$ . Overall, HI is  $\sim 5\%$  ( $18\%$ ) more abundant in F6 (F5) than in GR at  $z = 3$ . At  $z = 2$  there is  $\sim 12\%$  ( $22\%$ ) more HI in F6 (F5) than GR, while at  $z = 1$  we find more HI in F6 than in F5 and GR. This behavior can be understood as follows. At high redshifts ( $z \gtrsim 4-5$ ), modified gravity effects on the matter and halo distribution are screened for the models considered here, and thus F5 and F6 both behave similarly to GR. At intermediate redshifts ( $z = 2 \sim 3$ ), low-mass halos in F5 and F6 become unscreened and experience enhanced growth, leading to increased abundance of these low-mass objects compared to GR. Since neutral hydrogen can survive only in self-shielding halos, this implies that in F5 and F6 there are more HI-hosting halos than in GR and, consequently, these models are characterized by larger overall HI abundance. At low redshifts, baryonic effects become important. At these redshifts, we suspect that processes of gas heating can be more efficient in MG than in GR, reducing the overall HI abundance in F5 and F6 compared to GR. A closer inspection of Fig. 1 suggests that F6 behaves as a “retarded” version of F5, with the maximum enhancement with respect to GR shifted to lower redshifts. This is expected as the screening is more efficient in F6 and thus leads to a later onset of the MG force enhancement compared to F5.

## B. HI mass in halos

In this subsection, we present and discuss the *halo HI mass function*, i.e., the average HI mass enclosed in halos as a function of the halo mass. In the postreionization epoch, the majority of HI resides in halos, and therefore an accurate knowledge of the halo HI mass function can be used to predict the HI power spectrum in real and redshift space without requiring the full hydrodynamical simulation apparatus, but instead by painting the HI on top of dark matter halos from DMO simulations [64]. This approximation has been applied to extract information on the HI power spectrum for nonstandard cosmological models of dark matter and dark energy [35–37].

Here, for each halo in our simulations we measure its enclosed HI mass. The halos are identified using the SUBFIND [65] algorithm implemented in AREPO. The halo mass  $M_{\text{halo}}$  is defined as  $M_{200}$ , the mass contained in a sphere of radius  $r_{200}$ , within which the average density is 200 times the critical density at the specified redshift. In Fig. 2, we plot the HI mass as a function of the host halo mass for halos in the S25 simulations. As the figure shows, the HI mass increases monotonically with the halo mass on average in all models considered in our analysis, which is in agreement with what was found in [34]. The plot also shows that HI is present within very low-mass halos,

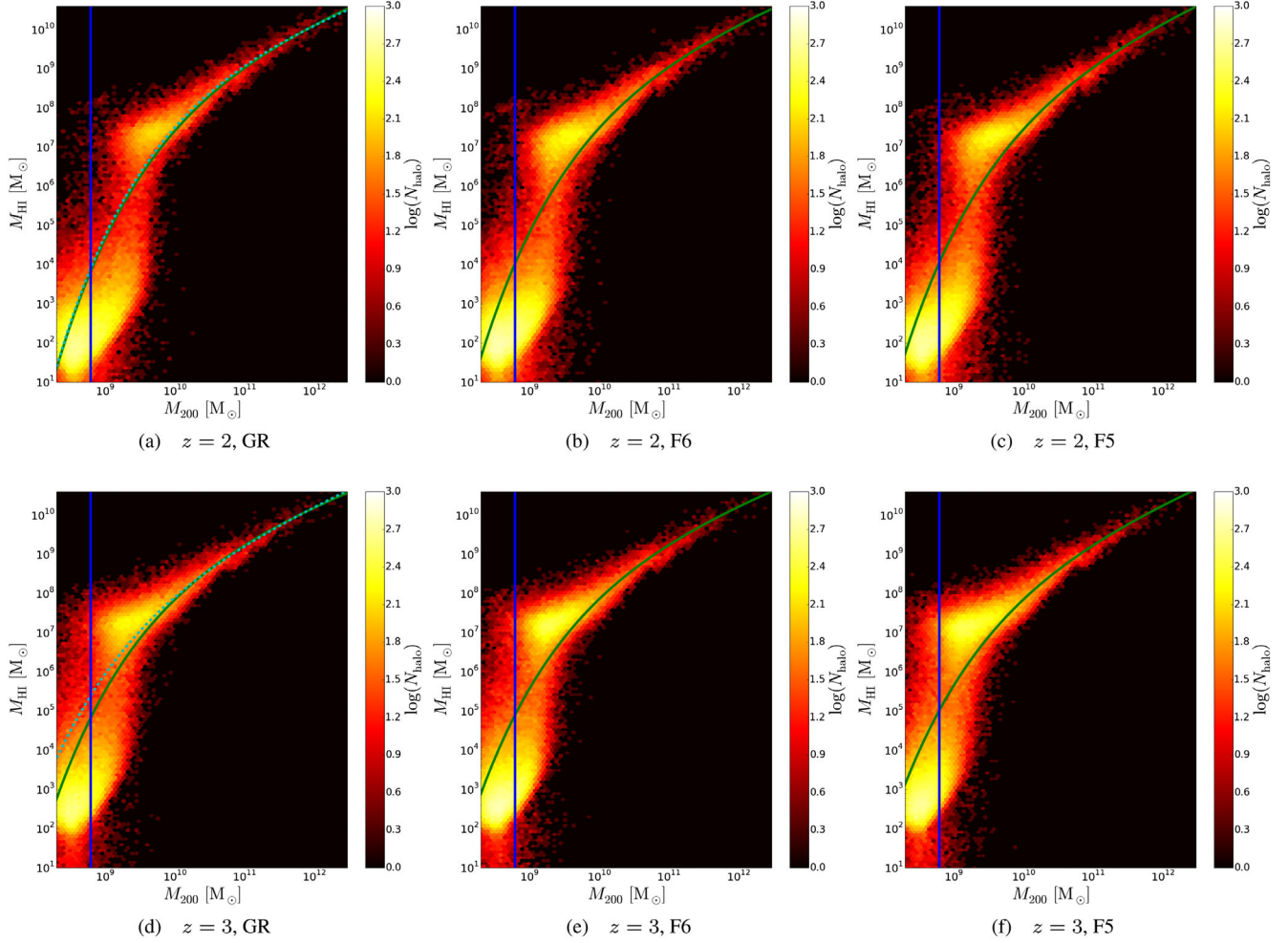


FIG. 2. HI mass contained in each halo at  $z = 3$  and  $z = 2$  measured from the S25 simulations for GR, F6, and F5, respectively (as labeled). The green solid line in each panel shows the best-fit curve obtained from Eq. (9). The best-fit values of the free parameters are displayed in Table I. The cyan dotted line indicates the best-fit curve of Eq. (9) found in [34] for GR. The blue vertical line indicates the mass of halos with around 50 DM (simulation) particles in S25.

$\sim 10^9 M_\odot$ , which highlights the importance of the simulation resolution for HI clustering estimates.

In [34], the authors identified a fitting function,

$$M_{\text{HI}}(M_{200}, z) = M_0 \left( \frac{M_{200}}{M_{\text{min}}} \right)^\alpha \exp \left[ - \left( \frac{M_{\text{min}}}{M_{200}} \right)^{0.35} \right], \quad (9)$$

which captures the power-law behavior at high halo masses,  $M_{\text{HI}} \propto M_{200}^\alpha$ , and the exponential cutoff for masses  $M_{200} \lesssim M_{\text{min}}$ . We fit our data with this formula by dividing the halo mass range  $\log(M_{200}/M_\odot) \in [8.0, 12.5]$  into two subranges:  $[8.0, 11.0]$  and  $[11.0, 12.5]$ . In order to account for the fact that at larger masses the halo sample is much smaller, we divide these two subranges into 40 and 9 bins, respectively, calculate the mean and variance of  $\log(M_{\text{HI}}/M_\odot)$  in each bin, and use a minimum- $\chi^2$  method to obtain the best-fit parameters. We find that this function remains a good fit to our simulations for all the models studied here (see Fig. 2). The best-fit values for the free

parameters  $\{M_0, \alpha, M_{\text{min}}\}$  are shown in Table I for  $z = 2$  and  $z = 3$  and for GR, F6, and F5.

In Fig. 2, we also compare our best-fit curves with the GR results taken from [34] (cyan dotted lines). As we can see, the fitting results in the two works agree very well at  $z = 2$  for the entire halo mass range. At  $z = 3$ , our result is again in good agreement with [34] for  $M_{200} \gtrsim 5 \times 10^9 M_\odot$ ,

TABLE I. Best-fit values of the free parameters for Eq. (9) at  $z = 2$  and 3 for GR, F6, and F5 (see main text for details).

Model	$z$	$M_0 [M_\odot]$	$\alpha$	$M_{\text{min}} [M_\odot]$
GR	3.0	$1.30 \times 10^{10}$	0.62	$2.72 \times 10^{11}$
F6	3.0	$1.19 \times 10^{10}$	0.65	$2.40 \times 10^{11}$
F5	3.0	$8.81 \times 10^9$	0.72	$1.79 \times 10^{11}$
GR	2.0	$1.69 \times 10^{10}$	0.67	$4.80 \times 10^{11}$
F6	2.0	$1.52 \times 10^{10}$	0.66	$4.25 \times 10^{11}$
F5	2.0	$1.34 \times 10^{10}$	0.73	$3.65 \times 10^{11}$

while slightly disagreeing at lower masses. This can be due to differences in the details of the fitting procedure or due to cosmic variance because of our smaller simulation box.

A quick comparison of our results with Fig. 4 in [34] shows that, again because of the smaller box size, our high-resolution simulations do not contain halos more massive than  $\sim 10^{12} M_\odot$ . This can explain why our total HI abundance (see III A) is slightly lower than that found in [34]. It is nevertheless possible to correct the total HI abundance in our simulations for the missing contribution from high-mass halos,  $\Omega_{\text{HI}}^{\text{corr}}$ . Given the expression for the total  $\Omega_{\text{HI}}^{\text{tot}}$  in the Universe (and assuming that all HI is contained in halos):

$$\Omega_{\text{HI}}^{\text{tot}} = \frac{1}{\rho_{\text{c0}}} \int_0^\infty M_{\text{HI}}(M_{200}) \frac{dn_{\text{halo}}}{dM_{200}} dM_{200}, \quad (10)$$

in which  $dn_{\text{halo}}/dM_{200}$  is the halo mass function (HMF) and  $M_{\text{HI}}(M_{200})$  is given by Eq. (9), the missing HI in the simulations can be estimated as

$$\Omega_{\text{HI}}^{\text{corr}} = \frac{1}{\rho_{\text{c0}}} \int_{M_{200}^{\text{cut}}}^\infty M_{\text{HI}}(M_{200}) \frac{dn_{\text{halo}}}{dM_{200}} dM_{200}. \quad (11)$$

Here  $M_{200}^{\text{cut}}$  is the maximum halo mass in the considered simulation, while the HMF is estimated using the approach of [66]. Equation (11) gives an approximate estimate for the HI fraction in large halos which are missing in S25,  $\Omega_{\text{HI}}^{\text{corr}} \approx 7 \times 10^{-5}$  at  $z = 3$  for GR. Adding this contribution to the  $\Omega_{\text{HI}}$  measured from our simulations (cf. Fig. 1) brings the total  $\Omega_{\text{HI}}$  in very good agreement with that found in [34].

As mentioned already, the above fitting results for the halo HI mass function can be used to model the HI distribution in DMO simulations, without having to run computationally expensive high-resolution hydrodynamical simulations, but using a *halo occupation distribution* technique applied to HI (see, e.g., [64] for more details).

### C. HI clustering

In the previous subsections, we have focused on the overall HI abundance and the halo HI mass function. Although these quantities provide useful information about the total HI in our simulation boxes and the HI inside halos, they do not directly describe the HI distribution and clustering. To understand the differences in the matter clustering of HI, we now analyze the HI power spectrum,  $P_{\text{HI}}(k)$ .

We calculate  $P_{\text{HI}}(k)$  in redshift space from the simulation outputs by applying a cloud-in-cell density assignment scheme ([67,68]), while the HI power spectra in real space from [46] (these results were obtained from the same simulation suite) are shown for comparison. In the left and central panels of Fig. 3, we display the measured power spectra for the HI overdensity,  $\rho_{\text{HI}}/\bar{\rho}_{\text{HI}}$ , where  $\bar{\rho}_{\text{HI}}$  is the mean HI density defined above. We present the  $P_{\text{HI}}(k)$

results measured from both the S62 (dashed lines) and S25 (solid lines) simulations. The overall effect in F6 and F5 (seen in both the low- and high-resolution simulations and in both real and redshift space) is a scale-dependent reduction of the clustering power with respect to GR. However, the results for the high- and low-resolution simulations do not converge at the scales probed by the simulations. As mentioned above, the disagreement in the results from S62 and S25 reflects the lack of resolved small halos with  $M_{200} < 10^{10} M_\odot$  in the former. For this reason, we will only discuss the results from the S25 simulations in the following.

For the monopole of the redshift space power spectrum, we find that at  $z = 3$  the F6 (F5)  $P_{\text{HI,redshift}}(k)$  is suppressed by 8% (14%) for  $k \sim 2h \text{ Mpc}^{-1}$ , while the effect is even stronger at higher wave numbers. At  $z = 2$ , for F6 (F5) it is suppressed by 13% (18%) with respect to GR for  $k \sim 2h \text{ Mpc}^{-1}$ . Similar trends can be found for the real-space power spectrum. It is interesting to note that at  $z = 2$  the F6  $P_{\text{HI,real}}(k)$  is slightly more suppressed than that for F5 at all scales probed by our simulations, while at  $z = 3$  the suppression is stronger for F5.

To understand the above results, in the right panels of Fig. 3, we compare to the HMFs of the SHYBONE simulations (taken from [69]). In the case of S62, we show the HMFs for both the full-physics (dashed lines) and DMO (symbols) simulations, while for S25 we only show the full-physics ones (solid lines). The ratios relative to GR for F5 and F6 measured from full-physics S62 and full-physics S25 agree very well for halos with masses  $> 3 \times 10^{10} M_\odot$ . However, due to the lower resolution of S62 (halos with  $M_{200} \sim 10^{10} M_\odot$  contain roughly 50 DM particles in this box), the HMFs disagree at lower masses. Since halos with  $10^9 M_\odot < M_{200} < 10^{10} M_\odot$  can host appreciable amounts of HI [34], this explains why the ratios of the  $f(R)$  HI power spectra with respect to GR for the large box do not agree with those measured from the small box.

Analyzing the behavior of the HMF ratios, we find that at  $z \geq 2$  F5 and F6 are characterized by a larger number of low-mass halos ( $M_{200} \lesssim 10^{12} M_\odot$ ) than GR. As HI can survive only in halos where self-shielding effects prevent it from ionization, in MG there are more hosts for HI than GR. Therefore, our interpretation of the behavior of the HI power spectra in the different models is that it primarily reflects the differences in the HMFs of these models (though these models also have different halo density profiles [70], which can have impact on the HI distribution as well), with  $f(R)$  gravity being able to turn more low initial density peaks into halos. Given that low initial density peaks are less clustered, HI, as a tracer of them, has a smaller clustering amplitude in  $f(R)$  models compared to GR.

Comparing the HMFs measured from the DMO and the full-physics simulations in the right column of Fig. 3, it is obvious that the galaxy formation processes have a



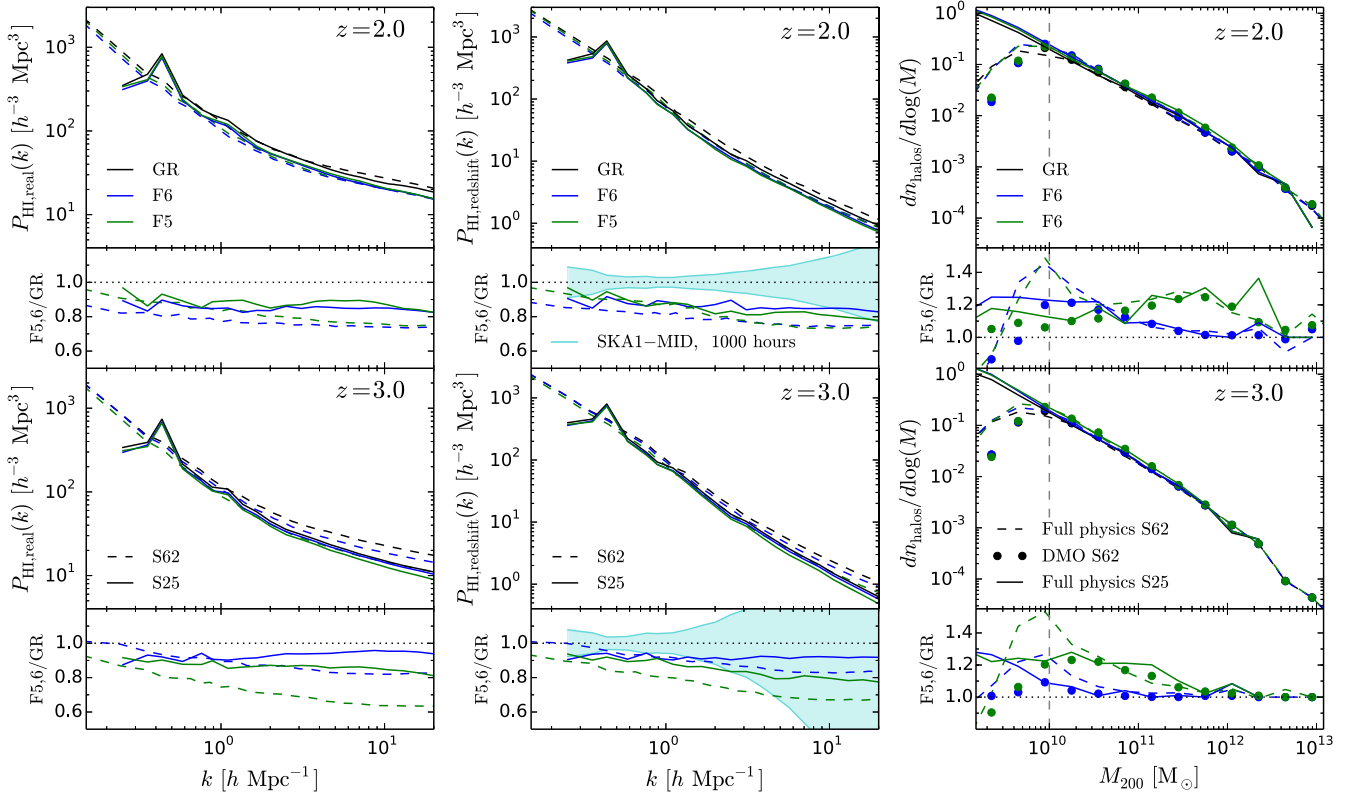


FIG. 3. Neutral hydrogen power spectra and halo mass functions at  $z = 2$  (upper panels) and  $z = 3$  (lower panels) for GR (black), F6 (blue), and F5 (green). Left panels: Real-space HI power spectra  $P_{\text{HI}}$ . Dashed lines show the results for S62, while solid lines show those for S25. The real-space power spectra are taken from Ref. [46]. Central panels: Same as the left panels but for the redshift space (monopole) HI power spectra. The cyan shaded areas in the lower subpanels represent the expected errors for SKA1-MID measurements for GR, assuming 1000 observing hours. The error bars are calculated using the GR redshift space power spectrum measured from S25. Right panels: Differential halo mass functions as shown in [69], for comparison. Solid (dashed) lines show the results from full-physics S25 (S62) simulations. Symbols indicate the results from DMO S62 simulations. The lower subpanels show the relative differences to GR.

non-negligible effect on the HMF itself. The relative differences in the HMFs between the  $f(R)$  models and GR (lower panels in the right column) are nevertheless only mildly affected by the baryons for  $M_{200} \gtrsim 3 \times 10^{10} M_{\odot}$  at  $z = 2$  and 3. At lower masses, we suspect that the differences between the full physics and DMO results are due to resolution effects.

#### D. Additional tests

To further check the above result about the different behavior of  $P_{\text{HI}}$  in MG and GR, we have carried out two additional tests.

Figure 4 compares the HI power spectrum (the *actual*  $P_{\text{HI}}$ ) in real space with a *halo* HI power spectrum,  $P_{\text{HI,halo}}$ , calculated by assuming that for each halo all the HI contained in it is at its center. As shown in this figure, the ratios of  $P_{\text{HI,halo}}$  for  $k < 2h \text{ Mpc}^{-1}$  are very similar (within a few %) to those from the *actual*  $P_{\text{HI}}$ , confirming that the model differences in  $P_{\text{HI}}$  at large scales are determined by halo clustering. However, at even larger  $k$ , the  $P_{\text{HI,halo}}$  results start to deviate from the *actual*  $P_{\text{HI}}$

because the former do not account for the spatial distribution of HI inside the halos (see, e.g., Ref. [34]), and this affects the relative difference between the  $f(R)$  models and GR. Since the differences in the *actual* power spectrum among the models closely follow those in  $P_{\text{HI,halo}}$  at large scales, we can conclude that the HI distribution is more sensitive to the clustering of halos than to the effects of baryons at these scales. As we can see from Fig. 4, the above result is true for both simulation setups, including S62. Because in S62 we do not have accurate information on the clustering of halos with masses  $< 10^{10} M_{\odot}$ , inevitably the HI power spectrum results therein are not accurate and tend to deviate from their high-resolution counterparts.

To further show the importance of simulation resolution for accurate predictions of HI, Fig. 5 displays the number density of halos with HI mass  $M_{\text{HI}} \geq 10^6 M_{\odot}$  (i.e., the halos that contribute significantly to  $P_{\text{HI}}$ ). First, we can see that in the case of S25 the number of HI-rich halos increases with decreasing halo mass and peaks at a certain mass scale,  $M_{200}^{\text{peak}}$  (for GR,  $M_{200}^{\text{peak}} \sim 2 \times 10^9 M_{\odot}$  at  $z = 3$  while  $M_{200}^{\text{peak}} \sim 4 \times 10^9 M_{\odot}$  at  $z = 2$ ). For masses lower



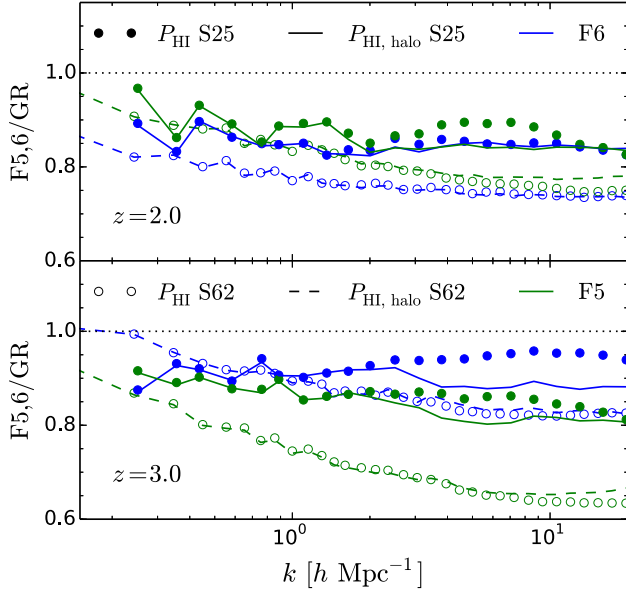


FIG. 4. The ratios of the F5 (green) and F6 (blue) HI power spectra with respect to GR for the *actual* real-space HI power spectrum (symbols) and the *halo* HI power spectrum (lines),  $P_{HI, halo}$ , as defined in the text, at  $z = 2.0$  (upper panel) and  $z = 3.0$  (lower panel). Solid lines and full circles show the results from S25, while dashed lines and empty circles show the results from S62.

than  $M_{200}^{\text{peak}}$ , the number of HI-rich halos decreases, though the halo mass function keeps increasing. Because in S25 halos with 50 DM particles have a mass of  $\sim 6 \times 10^8 M_\odot$ , the halos near (including a factor of a few below)  $M_{200}^{\text{peak}}$  are

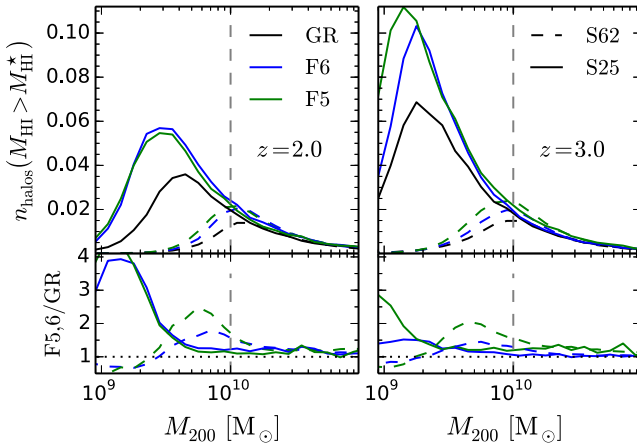


FIG. 5. Number density of halos with HI mass  $\geq M_{HI}^* = 10^6 M_\odot$ , for GR (black), F6 (blue), and F5 (green) at  $z = 2.0$  (left panel) and  $3.0$  (right panel). Solid lines show the results for S25, while dashed lines represent those from S62. The vertical grey dashed lines indicate halos with  $\sim 50$  particles in S62. We do not show the corresponding limit for S25 because it is at around  $\sim 6 \times 10^8 M_\odot$ , so it is outside the halo mass range shown in the figure.

well resolved. This is not the case for S62, for which the peak mass  $M_{200}^{\text{peak}}$  predicted by simulations coincides with the resolution limit ( $M_{200}^{\text{peak}} \sim 10^{10} M_\odot$ , see the vertical dashed line in Fig. 5) for the simulation, implying that the decrease in the HI-rich halo abundance at  $M_{200} < M_{200}^{\text{peak}}$  is probably a resolution effect. We therefore conclude that the S25 simulations have a sufficient resolution for predicting the HI abundance and clustering.

Regarding the overall low-mass HI-rich halo abundance for S25, we find that at  $z = 3$  it is higher in F5, F6 than in GR, suggesting that HI is distributed in a less clustered way in the MG models. At  $z = 2$ , the number of HI-rich halos in F6 is slightly larger than that in F5 for halo masses  $\gtrsim 2 \times 10^9 M_\odot$ , which may explain why the HI power spectrum is a little more suppressed in F6 than in F5 at this redshift.

Another interesting observation from Fig. 5 (see the results for S25 in the low panel) is that the HI-rich halo abundances for F6 and F5 are close to each other at  $z = 2$  and  $z = 3$ . This can explain why we find a very similar degree of suppression in the power spectra of these two models in the previous subsection. This behavior of the HI-rich halo mass function can be understood as follows. Because of its less efficient chameleon screening, one may naively expect that F5 is able to turn more initial density peaks into halos (above a certain mass) than F6, and so this model always has more HI-rich halos than F6 at a given time. However, continuously decreasing the efficiency of screening cannot increase the number of low-mass halos indefinitely due to increased merger rates and because the maximum enhancement of the strength of gravity is  $4/3$  in  $f(R)$  gravity. Furthermore, the overall HI available to form inside halos is also limited. Figure 5 implies that there is an upper limit on the number of HI-hosting halos that can be produced and making structures unscreened earlier will not help to increase the HI halo abundance beyond this “saturation point.” This is particularly evident from the results at  $z = 2$ , where F5 and F6 have nearly identical distributions. This result is interesting since two models with very different screening strengths, such as F6 and F5, can end up having the same constraint.

### E. Observational forecast

To understand the extent to which future measurements of  $P_{HI}$  in redshift space can distinguish F6 and F5 from GR, we estimate the  $1\sigma$  errors on the power spectrum expected from the instrumental noise of SKA1-MID radio telescope [25] for GR, following the method given in [35–37, 71] and using a realistic baseline densities computed in [32]. We compute the expected  $1\sigma$  errors for 1000 observing hours, as shown by the shaded areas in the central lower subpanels of Fig. 3. In the calculation, we have used the monopole of the HI redshift space power spectrum for GR measured from the S25 simulation. Comparing the errors with the

TABLE II. The integrated S/N ratios for distinguishing a MG model from GR using redshift space  $P_{\text{HI}}(k)$  with  $k_{\text{max}} = 2h \text{ Mpc}^{-1}$ , with  $(\text{S/N})^2 \equiv \sum_{k_{\text{min}}}^{k_{\text{max}}} [P_{\text{HI}}^{\text{MG}}(k) - P_{\text{HI}}^{\text{GR}}(k)]^2 / \sigma^2(k)$ ; here  $\sigma(k)$  is the expected  $1\sigma$  error from SKA1-MID for 1000 hours, while  $k_{\text{min}}$  is set by the value of the box length  $L = 25h^{-1} \text{ Mpc}$  of our high-resolution simulations,  $k_{\text{min}} = 2\pi/L$ . The above results are calculated using the redshift space HI power spectra measured from S25 only.

S/N	$z = 3.0$	$z = 2.0$
F6	4.3	10.1
F5	5.5	9.9

$P_{\text{HI}}$  ratios with respect to GR in redshift space for S25 (solid lines), we find that F6 and F5 can be both distinguished from GR at  $z = 2$  and for  $k < 2h \text{ Mpc}^{-1}$  by using a 1000-hour integration. The integrated signal-to-noise (S/N) ratios for distinguishing MG to GR are shown in Table II for the two redshifts considered in this analysis.

### F. Uncertainties in subgrid physics

Current state-of-the-art simulations are still unable to fully resolve the formation and evolution of stars and galaxies from first principles, largely due to the huge dynamical range between the scales at which star formation and black hole accretion take place and the scales which must be covered in order to realistically reproduce the large-scale structure of the Universe. As a result, these simulations rely on simplified models to approximate the main small-scale baryonic processes, such as stellar and black hole evolution and their feedback, beyond the resolution limits. These are referred to as “subgrid” recipes—such as the IllustrisTNG galaxy formation model used in this work—and usually include a number of free parameters that are calibrated against a set of observational data.

The IllustrisTNG model was calibrated against the stellar mass function, stellar mass fraction, galaxy central black hole masses and gas fraction, galaxy sizes at redshift  $z = 0$ , and the cosmic star formation rate history. Not surprisingly, the approximate nature of the subgrid physics can introduce uncertainties and degeneracies in the results. Two of the questions, which follow immediately, are thus: (1) can the IllustrisTNG subgrid model, which is calibrated for  $\Lambda\text{CDM}$ , be used in the  $f(R)$  simulations without any changes, and (2) does the simplified implementation of baryonic physics severely affect the signal expected from modified gravity? The first question was addressed in Ref. [46], where it was found that a recalibration is not necessary, because of the large uncertainties in observational data and the relatively modest effect of modified gravity on the observables used for calibration. Concerning the second question, we do expect the HI distribution to depend on the underlying galaxy formation model.

Unfortunately, there is no easy way to quantify the effect of changing the model without running a large number of simulations with different subgrid physics. This latter effort is beyond the scope of this work and would still leave plenty room for uncertainties. We will therefore limit our discussion below to qualitative comments on the possible effects of changing subgrid physics on our results.

An important point to understand is to what extent the uncertainties in active galactic nuclei (AGN) and stellar feedback can affect the results above. We consider first AGN feedback, which can expel HI from the centers of massive halos or ionize the gas [34], hence affecting its clustering. While the impact of AGN feedback on the total matter power spectrum  $P_{\text{tot}}(k)$  at  $z = 0$  is found to be strong in previous works (see, e.g., [72]), more recent high resolution simulations of galaxy formation, such as EAGLE (see Fig. 1 in [73]) and IllustrisTNG (see Fig. 7 in [49]), agree remarkably well and find it to be less than 1% on scales  $k \lesssim 2h \text{ Mpc}^{-1}$ . At higher redshifts, which we are interested in here, the effect is even weaker.

Moreover, at scales  $k < 2h \text{ Mpc}^{-1}$  (cf. Sec. III D) the MG effects on the HI clustering are mainly due to changes in low-mass ( $< 10^{10} M_{\odot}$ ) unscreened objects which do not host AGN. More massive, AGN-hosting halos tend to be screened, particularly at high redshift. The halo mass at which chameleon screening becomes efficient can be estimated using the fitting formula presented in Ref. [74], and we find that halos more massive than  $2 \times 10^9 M_{\odot}$  and  $2.5 \times 10^{10} M_{\odot}$  in F6 are screened at  $z = 3$  and  $z = 2$ , respectively, while AGN feedback mainly influence halos of  $M_{200} > 4 \times 10^{12} M_{\odot}$  [34]. As a result, AGN feedback and modified gravity affect two different regimes of halo masses.

We do not expect moderate changes in stellar feedback to have a strong impact on the relative difference between the HI power spectra of the models either. This is because at large scales ( $k < 2h \text{ Mpc}^{-1}$ ) the clustering of HI closely mimics that of the dark matter halos which host HI, while stellar feedback mainly affects the distribution of gas and stars within the halos and thus smaller scales. Therefore, at least for the feedback prescription used in IllustrisTNG, stellar feedback does not seem to significantly change the key signature of modified gravity on dark (and total) matter clustering. Much stronger variations to the feedback mechanism may alter the HI content within halos of a given mass range and consequently the HI power spectrum; such strong changes would nevertheless lead to tensions with the low redshift observables used to tune the IllustrisTNG model [53]. We also note that in Ref. [75] the authors analyzed the impact of different physical processes on the HI column density, and found that a significant effect (note that the variations of the subgrid physics parameters in that work are extreme) is only observed in the regime of very high column density, corresponding to the very inner regions of massive halos where the AGN feedback is strong.

From the discussion above, we expect the main conclusions of this work to be relatively robust against changes of a subgrid physics parameter in the simulations. This said, given that HI clustering is a promising probe of cosmology for the future, it would be useful to study in greater detail how it is affected by the uncertain subgrid physics associated with galaxy formation—not only in the context of modified gravity, but also for the  $\Lambda$ CDM model itself.

#### IV. SUMMARY AND CONCLUSIONS

The screening mechanism of chameleon-type MG models, such as  $f(R)$  gravity, is particularly efficient at high redshift and for massive objects. Consequently, the cosmological probes proposed to date focus primarily on low redshift. The constraining power, mainly from high-mass objects, is nevertheless still limited due to the efficient screening in these objects, while it is difficult to obtain accurate cosmological data for less-screened, low-mass objects.

In this paper, we explore a different approach to constrain  $f(R)$  gravity by making use of the fact that the low-mass end of the halo mass function is enhanced in this MG model already at redshift  $z \sim 2-3$ . We propose that this enhancement should be observable through 21-cm intensity mapping and use the SHYBONE simulations, a set of state-of-the-art hydrodynamical simulations employing the IllustrisTNG model which were carried out for two different  $f(R)$  gravity models (F5 and F6), to analyze the viability of this approach. Our results can be summarized as follows.

- (i) The ratios of the overall HI abundance with respect to GR follow a similar trend for F6 and F5. Both models predict similar  $\Omega_{\text{HI}}$  as GR at  $z \gtrsim 4$ . At lower redshifts, the ratios increase, reaching a maximum enhancement at  $z = 2$  and  $z = 1$  for F6 and F5, respectively, before starting to decrease at even lower redshifts. The results from S62 are limited by the resolution, since these simulations do not resolve halos with masses  $M_{200} \lesssim 10^{10} M_{\odot}$ , which contain non-negligible amounts of HI.
- (ii) The HI mass enclosed in halos for GR, F6, and F5 can be well described by the fitting formula proposed in [34], which approaches a power law at high halo mass, and is exponentially suppressed at low masses. We provide the best-fit values of the parameters of this formula for all the models considered here at  $z = 2$  and 3. Our fitting curves for GR are in agreement with those found in [34]. These results can be used to model the HI distribution and clustering in  $f(R)$  gravity using DMO simulations by painting the HI onto the center of each dark matter halo.
- (iii) The HI power spectrum (both in real and redshift space) in F5 and F6 is suppressed with respect to that measured for GR. This suppression can be detected

in future 21-cm experiments such as SKA1-MID with 1000 hours of exposure time at a S/N ratio of  $\sim 10$  at  $z = 2$  for both F5 and F6.

- (iv) The differences in the HI power spectrum closely reflect the differences in the HI-hosting halo abundance among the models. The halo mass functions in F5 and F6 are enhanced with respect to GR at  $M_{200} \lesssim 10^{12} M_{\odot}$ , showing that  $f(R)$  gravity is able to turn more low-density peaks into halos which host HI. Since low density peaks are less clustered, HI power spectra have lower amplitudes in MG compared to GR. The above statement is corroborated by the analysis of the HI-hosting halo power spectrum, i.e., the power spectrum calculated assuming the HI in each halo is concentrated at the center of that halo. This test shows that the HI power spectrum for  $k < 2h \text{ Mpc}^{-1}$  is influenced more by halo distribution and clustering rather than by baryonic effects.
- (v) The predicted HI distribution strongly depends on the resolution of the simulations. Indeed, as can be seen from the abundance of HI-rich halos (Fig. 5), the S62 simulations are not able to resolve halos with  $M_{200} \lesssim 10^{10} M_{\odot}$ ; consequently we have no information on the HI inside such halos and these simulations are not able to give accurate predictions of the relative differences in the HI power spectra for F6 and F5 with respect to GR. The S25 simulations are, in contrast, able to resolve halos with  $M_{200} \gtrsim 10^9 M_{\odot}$ , predicting more HI-hosting halos than in the case of their low-resolution counterparts.

Our results indicate that 21-cm intensity mapping can prove useful in constraining  $f(R)$  gravity models. For F6, for example, future 21-cm intensity experiments can offer a strong test compared with the other cosmological probes which focus on low-redshifts and massive objects. While this seems promising, however, we note that the HI distribution may depend on the underlying galaxy formation model chosen for the simulations. In Sec. III F, we discuss the main processes (AGN and stellar feedback) which can affect the HI clustering. However, a comprehensive study about how the different baryonic effects can affect the HI in galaxies does not exist yet, and would be an interesting topic for future work.

As a side remark, we note that 21-cm intensity mapping can also give accurate information of the expansion history [76–78], which may be used to break potential degeneracy between modified expansion history and structure growth. This is because viable  $f(R)$  models with a working chameleon screening mechanism to restore GR in the Solar System must have practically identical expansion history to  $\Lambda$ CDM, but structures in these models still grow differently at high  $z$ .

To conclude, 21-cm intensity mapping represents a new and potentially useful tool to constrain nonstandard



cosmological models that modify the matter and halo distribution from predictions by  $\Lambda$ CDM. The results of this work suggest that the HI distribution is appreciably affected by MG effects and can be used to shed light on the nature of gravity itself.

### ACKNOWLEDGMENTS

We thank Carlton Baugh, Jianhua He, Tom Theuns, and Francisco Villaescusa-Navarro for useful discussions. The authors are supported by the European Research Council via Grant No. ERC-StG-716532-PUNCA. B. L. is additionally supported by Science and Technology Facilities

Council (STFC) Consolidated Grant No. ST/P000541/1. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC Distributed Research using Advanced Computing (DiRAC) High Performance Computing (HPC) Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). The equipment was funded by Department for Business, Energy and Industrial Strategy (BEIS) capital funding via STFC capital Grants No. ST/K00042X/1, No. ST/P002293/1, No. ST/R002371/1, and No. ST/S002502/1, Durham University and STFC operations Grant No. ST/R000832/1. DiRAC is part of the National e-Infrastructure.

- 
- [1] C. M. Will, *Living Rev. Relativity* **17**, 4 (2014).
  - [2] K. Koyama, *Rep. Prog. Phys.* **79**, 046902 (2016).
  - [3] M. Cataneo, D. Rapetti, F. Schmidt, A. B. Mantz, S. W. Allen, D. E. Applegate, P. L. Kelly, A. von der Linden, and R. G. Morris, *Phys. Rev. D* **92**, 044009 (2015).
  - [4] H. Wilcox *et al.*, *Mon. Not. R. Astron. Soc.* **452**, 1171 (2015).
  - [5] X. Liu *et al.*, *Phys. Rev. Lett.* **117**, 051101 (2016).
  - [6] S. Peirone, M. Raveri, M. Viel, S. Borgani, and S. Ansoldi, *Phys. Rev. D* **95**, 023521 (2017).
  - [7] L. Lombriser, F. Simpson, and A. Mead, *Phys. Rev. Lett.* **114**, 251101 (2015).
  - [8] Y. Cai, N. Padilla, and B. Li, *Mon. Not. R. Astron. Soc.* **451**, 1036 (2015).
  - [9] M. Cautun, E. Paillas, Y.-C. Cai, S. Bose, J. Armijo, B. Li, and N. Padilla, *Mon. Not. R. Astron. Soc.* **476**, 3195 (2018).
  - [10] J. Armijo, Y.-C. Cai, N. Padilla, B. Li, and J. A. Peacock, *Mon. Not. R. Astron. Soc.* **478**, 3627 (2018).
  - [11] C. Hernandez-Aguayo, C. M. Baugh, and B. Li, *Mon. Not. R. Astron. Soc.* **479**, 4824 (2018).
  - [12] A. Terukina, L. Lombriser, K. Yamamoto, D. Bacon, K. Koyama, and R. C. Nichol, *J. Cosmol. Astropart. Phys.* **04** (2014) 013.
  - [13] H. A. Buchdahl, *Mon. Not. R. Astron. Soc.* **150**, 1 (1970).
  - [14] P. Brax, A.-C. Davis, B. Li, and H. A. Winther, *Phys. Rev. D* **86**, 044015 (2012).
  - [15] W. Hu and I. Sawicki, *Phys. Rev. D* **76**, 064004 (2007).
  - [16] J. Khoury and A. Weltman, *Phys. Rev. D* **69**, 044026 (2004).
  - [17] J. Khoury and A. Weltman, *Phys. Rev. Lett.* **93**, 171104 (2004).
  - [18] F. Schmidt, *Phys. Rev. D* **81**, 103002 (2010).
  - [19] G.-B. Zhao, B. Li, and K. Koyama, *Phys. Rev. Lett.* **107**, 071303 (2011).
  - [20] Y. Li and W. Hu, *Phys. Rev. D* **84**, 084033 (2011).
  - [21] L. Lombriser, B. Li, K. Koyama, and G.-B. Zhao, *Phys. Rev. D* **87**, 123511 (2013).
  - [22] E. Puchwein, M. Baldi, and V. Springel, *Mon. Not. R. Astron. Soc.* **436**, 348 (2013).
  - [23] C. Arnold, E. Puchwein, and V. Springel, *Mon. Not. R. Astron. Soc.* **440**, 833 (2014).
  - [24] C. Arnold, P. Fosalba, V. Springel, E. Puchwein, and L. Blot, *Mon. Not. R. Astron. Soc.* **483**, 790 (2019).
  - [25] P. E. Dewdney *et al.*, SKA Organisation, 2013, <http://skatelescope.org>.
  - [26] M. G. Santos *et al.* (MeerKLASS Collaboration), [arXiv: 1709.06099](https://arxiv.org/abs/1709.06099).
  - [27] M. P. van Haarlem *et al.*, *Astron. Astrophys.* **556**, A2 (2013).
  - [28] K. Bandura *et al.*, *Soc. Photo-Opt. Instrum. Eng. Conf. Ser.* **9145**, 914522 (2014).
  - [29] R. A. Battye, I. W. A. Browne, C. Dickinson, G. Heron, B. Maffei, and A. Pourtsidou, *Mon. Not. R. Astron. Soc.* **434**, 1239 (2013).
  - [30] S. Bharadwaj, B. B. Nath, and S. K. Sethi, *J. Astrophys. Astron.* **22**, 21 (2001).
  - [31] A. Loeb and J. S. Wyithe, *Phys. Rev. Lett.* **100**, 161301 (2008).
  - [32] P. Bull, P. G. Ferreira, P. Patel, and M. G. Santos, *Astrophys. J.* **803**, 21 (2015).
  - [33] M. G. Santos *et al.*, *Proc. Sci.*, AASKA14 (2015) 019.
  - [34] F. Villaescusa-Navarro *et al.*, *Astrophys. J.* **866**, 135 (2018).
  - [35] F. Villaescusa-Navarro, P. Bull, and M. Viel, *Astrophys. J.* **814**, 146 (2015).
  - [36] I. P. Carucci, F. Villaescusa-Navarro, M. Viel, and A. Lapi, *J. Cosmol. Astropart. Phys.* **07** (2015) 047.
  - [37] I. P. Carucci, P.-S. Corasaniti, and M. Viel, *J. Cosmol. Astropart. Phys.* **12** (2017) 018.
  - [38] E. McDonough and R. H. Brandenberger, *J. Cosmol. Astropart. Phys.* **02** (2013) 045.
  - [39] Y. Wang, Y. Xu, F. Wu, X. Chen, X. Wang, J. Kim, C. Park, K. G. Lee, and R. Cen, *Proc. Sci.*, AASKA14 (2015) 033.
  - [40] P. Brax, C. van de Bruck, A.-C. Davis, and D. J. Shaw, *Phys. Rev. D* **78**, 104021 (2008).
  - [41] J. Wang, L. Hui, and J. Khoury, *Phys. Rev. Lett.* **109**, 241301 (2012).
  - [42] J. J. Ceron-Hurtado, J.-h. He, and B. Li, *Phys. Rev. D* **94**, 064052 (2016).
  - [43] I. Sawicki and E. Bellini, *Phys. Rev. D* **92**, 084061 (2015).



- [44] B. P. Abbott *et al.* (LIGO Scientific, Virgo, Fermi-GBM, and INTEGRAL Collaborations), *Astrophys. J.* **848**, L13 (2017).
- [45] C. Burrage and J. Sakstein, *Living Rev. Relativity* **21**, 1 (2018).
- [46] C. Arnold, M. Leo, and B. Li, *Nat. Astron.* (2019), <https://dx.doi.org/10.1038/s41550-019-0823-y>.
- [47] V. Springel, *Mon. Not. R. Astron. Soc.* **401**, 791 (2010).
- [48] P. A. R. Ade *et al.* (Planck Collaboration), *Astron. Astrophys.* **594**, A13 (2016).
- [49] V. Springel *et al.*, *Mon. Not. R. Astron. Soc.* **475**, 676 (2018).
- [50] S. Genel *et al.*, *Mon. Not. R. Astron. Soc.* **474**, 3976 (2018).
- [51] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist, *Nature (London)* **509**, 177 (2014).
- [52] M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist, *Mon. Not. R. Astron. Soc.* **444**, 1518 (2014).
- [53] A. Pillepich *et al.*, *Mon. Not. R. Astron. Soc.* **475**, 648 (2018).
- [54] F. Marinacci *et al.*, *Mon. Not. R. Astron. Soc.* **480**, 5113 (2018).
- [55] D. Nelson *et al.*, *Mon. Not. R. Astron. Soc.* **475**, 624 (2018).
- [56] A. Pillepich *et al.*, *Mon. Not. R. Astron. Soc.* **473**, 4077 (2018).
- [57] R. Weinberger *et al.*, *Mon. Not. R. Astron. Soc.* **465**, 3291 (2017).
- [58] A. Rahmati, A. H. Pawlik, M. Raičević, and J. Schaye, *Mon. Not. R. Astron. Soc.* **430**, 2427 (2013).
- [59] S. M. Rao, D. A. Turnshek, and D. B. Nestor, *Astrophys. J.* **636**, 610 (2006).
- [60] P. Lah *et al.*, *Mon. Not. R. Astron. Soc.* **376**, 1357 (2007).
- [61] A. Songaila and L. L. Cowie, *Astrophys. J.* **721**, 1448 (2010).
- [62] P. Noterdaeme *et al.*, *Astron. Astrophys.* **547**, L1 (2012).
- [63] N. H. M. Crighton *et al.*, *Mon. Not. R. Astron. Soc.* **452**, 217 (2015).
- [64] F. Villaescusa-Navarro, M. Viel, K. K. Datta, and T. R. Choudhury, *J. Cosmol. Astropart. Phys.* **09** (2014) 050.
- [65] V. Springel, S. D. M. White, G. Tormen, and G. Kauffmann, *Mon. Not. R. Astron. Soc.* **328**, 726 (2001).
- [66] R. K. Sheth, H. J. Mo, and G. Tormen, *Mon. Not. R. Astron. Soc.* **323**, 1 (2001).
- [67] J. h. He, B. Li, and A. J. Hawken, *Phys. Rev. D* **92**, 103508 (2015).
- [68] Y. P. Jing, *Astrophys. J.* **620**, 559 (2005).
- [69] C. Arnold and B. Li, [arXiv:1907.02980](https://arxiv.org/abs/1907.02980).
- [70] C. Arnold, E. Puchwein, and V. Springel, *Mon. Not. R. Astron. Soc.* **448**, 2275 (2015).
- [71] I. P. Carucci, F. Villaescusa-Navarro, and M. Viel, *J. Cosmol. Astropart. Phys.* **04** (2017) 001.
- [72] M. P. van Daalen, J. Schaye, C. M. Booth, and C. D. Vecchia, *Mon. Not. R. Astron. Soc.* **415**, 3649 (2011).
- [73] W. A. Hellwing, M. Schaller, C. S. Frenk, T. Theuns, J. Schaye, R. G. Bower, and R. A. Crain, *Mon. Not. R. Astron. Soc.* **461**, L11 (2016).
- [74] M. A. Mitchell, J.-h. He, C. Arnold, and B. Li, *Mon. Not. R. Astron. Soc.* **477**, 1133 (2018).
- [75] G. Altay, T. Theuns, J. Schaye, C. M. Booth, and C. D. Vecchia, *Mon. Not. R. Astron. Soc.* **436**, 2689 (2013).
- [76] R. Ansari *et al.*, [arXiv:1810.09572](https://arxiv.org/abs/1810.09572).
- [77] F. B. Abdalla *et al.* (Cosmology SWG Collaboration), [arXiv:1501.04035](https://arxiv.org/abs/1501.04035).
- [78] P. Bull, *Astrophys. J.* **817**, 26 (2016).